Prepared from FORM PTO-1390	Transmittal Letter to the United States Designated/Elected Office (DO/EO/US)	JC10 Rec'd PCT/PTO 04 JAN 2002
Customer No.	026418	10/030062
Attorney's Docket No.:	GK-OEH-120 / 500814.20021	
U.S. Application No.:		
International Application No.:	PCT/DE00/02154	
International Filing Date:	JULY 4, 2000	4 JULY 2000
Priority Date Claimed:	JULY 5, 1999	5 JULY 1999
Title of Invention:	METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME	
Applicant(s) for (DO/EO/US):	Thomas MOORE and Anton HORN	

Applicant herewith submits to the United States Designated/Elected Office (DO/EO/US) the following items and other information:

[X] 1. This is a **FIRST** submission of items concerning a filing under 35 U.S.C. 371.
[ ] 2. This is a **SECOND** or **SUBSEQUENT** submission of items concerning a filing under 35 U.S.C. 371.
[ ] 3. This express request to begin national examination procedures [35 U.S.C. 371 (f)] at any time rather than delay examination until the expiration of the applicable time limit set forth in 35 U.S.C. 371(b) and PCT Articles 22 and
[ ] 4. A proper Demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.
[X] 5. A copy of Publication No. WO 01/02848 11JAN01 ~~the International Application as filed~~ [35 U.S.C. 371(c)(2)]
  a) __ is transmitted herewith (required only if not transmitted by the International Bureau)
  b) __ has been transmitted by the international Bureau
  c) __ is not required, as the application was filed in the United States Receiving Office (RO/US)
[X] 6. A translation of Publication No. WO 01/02848 11JAN01 ~~the International Application~~ into English [35 U.S.C. 371(c)(2)]
[ ] 7. Amendments to the claims of the International Application under PCT Article 19 [35 U.S.C. 371(c)(3)]
  a) __ are transmitted herewith (required only if not transmitted by the International Bureau)
  b) __ have been transmitted by the International Bureau
  c) __ have not been made; however, the time limit for making such amendments has **NOT** expired.
  d) __ have not been made and will not be made
[ ] 8. A translation of the amendments to the claims under PCT Article 19 [35 U.S.C. 371(c)(3)]
[X] 9. An UNSIGNED Oath or declaration of the inventor(s) [35 U.S.C. 371(c)(4)] *EXECUTED Decl/POA - TO FOLLOW*
[ ] 10. A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 [35 U.S.C. 371(c)(5)]

Items 11. to 16. Below concern other document(s) or information included:

[X] 11. An Information Disclosure Statement under 37 C.F.R. 1.97 and 1.98
[ ] 12. An Assignment document for recording. A separate cover sheet (PTO-1619A) in compliance with 37 CFR 3.28 and 3.31 is included.
[X] 13. _X_ A **FIRST** preliminary amendment
  __ A **SECOND** or **SUBSEQUENT** preliminary amendment
[X] 14. A **substitute specification and abstract (attached to the preliminary amendment)**
[ ] 15. A change of power of attorney and/or address letter
[X] 16. (other items or information) **PCT/RO/101, PCT/IB/332 26MAR01, PCT/IPEA/409 22AUG01, Search Report: (PCT/ISA/210) 8DEC00, PTO-1449 w/5 references.**

EXPRESS MAIL No.: **EL 915 668 453 US** Deposited: **January 4, 2002**
I hereby certify that this correspondence is being deposited with the United States Postal Service Express mail under 37 CFR 1.10 on the date indicated above and is addressed to: BOX PCT, Commissioner for Patents, Washington, DC 20231.

/Ruth Montalvo Date: January 4, 2002

U.S. Application No. (if known, see 37 C.F.R. 1.50):
International Application No.: **PCT/DE00/02154** Attorney's Docket No: **GK-OEH-120 / 500814.20021**

531 Rec'd PCT/ 04 JAN 2002

	CALCULATIONS	PTO USE ONLY
[X] 17. The following fees are submitted: BASIC NATIONAL FEE [37 CFR 1.492(a)(1)-(5)]		
[X] Search Report has been prepared by the EPO or JPO..... $ 890.00		
[ ] International preliminary examination fee paid to USPTO [37 CFR 1.482]..... $ 710.00		
[ ] No International preliminary examination fee paid to USPTO [37 CFR 1.482] but International search fee paid to USPTO [37 CFR 1.445(a)(2)..... $ 740.00		
[ ] Neither International preliminary examination fee [37 CFR 1.482] nor International search fee [37 CFR 1.445(a)(2)] paid to USPTO..... $ 1,040.00		
[ ] International preliminary examination fee paid to USPTO [37 CFR 1.482] and all claims satisfied provisions of PCT Article 33(1)-(4)..... $ 100.00		
ENTER APPROPRIATE BASIC FEE AMOUNT:	**$890.00**	

Claims	Number Filed		Number Extra	Rate		
Total Claims Prel. Amdt)	**12**	-20		x $ 18. =		
Indep. Claims	**1**	-03		x $ 80. =		
[__] Multiple Dependent Claim(s) (if applicable)				+ $ 270. =		

TOTAL OF ABOVE CALCULATIONS:	**$890.00**	
Surcharge of $130.00 for furnishing the oath or declaration later than [ ] 20 [ ] 30 months from the earliest claimed priority date [37 CFR 1.492(e)]		
TOTAL OF ABOVE CALCULATIONS:	**$890.00**	
Applicant claims Small Entity Status [See 37 CFR 1.27] Reduction by ½ for filing by small entity		
SUBTOTAL:	**$890.00**	
Processing fee of $130.00 for furnishing the English Translation later than [ ] 20 [ ] 30 months from the earliest claimed priority date [37 CFR 1.492(f)]		
TOTAL NATIONAL FEE:	**$890.00**	
Fee for recording the enclosed assignment [37 CFR 1.21(h)] The assignment must be accompanied by an appropriate cover sheet (PTO-1595) [37 CFR 3.28, 3.31].$ 40.00 per property +		
TOTAL FEE(S):	**$890.00**	
AMOUNTS TO BE REFUNDED OR CHARGED	**REFUNDED** $ **CHARGED** $	

(Please note the filing fee is based on the claims in the Preliminary Amendment)

[X] Check in the amount of **$ 890.00** to cover the above fees is enclosed. (The Commissioner is hereby authorized to charge any additional fees required with this submission or to credit any overpayment to Deposit Account No: 50-1529.)

NOTE: Where an appropriate time limit under 36 CFR 1.494 or 1.495 has not been met, a petition to revive [37 CFR 1.137(a) or (b)] must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:

Gerald H. Kiel, Esq. **(Customer No. 026418)**
Reed Smith LLP
375 Park Avenue
New York, NY 10152

Jules E. Goldberg — Name (Tel. (212) 521-5400)
Signature
24,408 — Reg. No.
January 4, 2002 — Date

EXPRESS MAIL mailing label No. **EL 915 668 453 US**
Date of Deposit: **January 4, 2002**
I hereby certify that this paper or fee is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" service under 37 CFR 1.10 on the date indicated above and is addressed to the Commissioner for Patents, Washington, D.C. 20231

Ruth Montalvo — 4 Jan 02 Date

Docket No.:GK-OEH-120/500814.20021

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicant:	Thomas MOORE and Anton HORN
Serial No.:	Unknown (Int'l Appln. PCT/DE00/02154 filed July 4, 2000)
Filed:	Simultaneously herewith
For:	METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME

PRELIMINARY AMENDMENT

Commissioner for Patents
Washington, D.C. 20231

Sir:

Please amend the above-identified application, filed simultaneously herewith, as follows:

IN THE SPECIFICATION

Cancel the present specification and substitute therefor the enclosed substitute specification.

IN THE CLAIMS

Cancel claims 1-12 and add new claims 13-24, reading as follows:

--13. (New) A method for the multidimensional analysis of a proteome in which the biological material with the proteome to be analyzed is solubilized and the proteins belonging to the proteome are separated, quantitatively determined and identified, comprising the steps of:

subjecting the proteome to a number n of different separating processes for n>2 under standardized conditions in such a way that each of the liquid fractions m_1 obtained in a separating step supplies m_2 liquid fractions in a subsequent separating steps, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions;

identifying said $m_1 * m_2 * \ldots m_n = M$ liquid fractions by τ different analysis processes qualitatively and/or quantitatively by identification processes, and determining said liquid ratio quantitatively by known quantification processes; and

after combining the analysis data, obtaining an n-dimensional image of the proteome which is characterized by identifiers and quantifiers and by the position in the n-dimensional data space.

14. (New) The method according to claim 13, wherein methods which separate according to the size of the protein and/or methods which separate according to the mass of the protein and/or methods which separate according to the charge of the protein and/or methods which separate according to the hydrophobicity of the protein and/or methods which separate according to the shape of the protein and/or methods which separate according to the affinity of the protein, with respect to specific ligands, also to antibodies are selected as separating methods.

15. (New). The method according to claim 13, wherein methods for determining specific immunological characteristics and/or methods for determining specific catalytic activity and/or methods for determining chemical

modification of the proteins of the proteome are used as identification methods.

16. (New). The method according to claim 13, wherein methods for nonspecific determination of protein concentration with different sensitivities and/or quantitative determination methods for determining specific catalytic activities and/or quantitative immunological methods and/or quantitative binding assays are selected as quantification methods.

17. (New). The method according to claim 13, wherein the identification of individual proteins of the proteome is carried out directly by mass determination of the proteins.

18. (New) The method according to claim 13, wherein the identification of individual proteins is carried out according to protease digestion and mass identification of fragments.

19. (New) The method according to claim 13 wherein, after the separation step, the fractions are assembled in a two-dimensional multiple vessel system, in the manner of and with the layout of microtitration plates.

20. (New). The method according to claim 13 wherein, in the first separating step, the fractions are assembled in a defined grid, preferably in the n * 96 grid of microtitration technology.

21. (New) The method according to claim 13, wherein all identification and quantification steps are carried out in a defined grid, preferably in the n * 96 grid, with adaptable liquid handling technique.

22. (New) The method according to claim 21, wherein all identification steps and quantification steps are carried out with at least four two-dimensionally arranged, simultaneously working pipettor channels.

23. (New). The method according to claim 13, wherein the first

dimension for separation is high-resolution size exclusion, ion exchange or hydrophobicity chromatography, which are known per se, in that the second dimension is carried out by parallel separation and fractionation of the fractions of the first dimension by a principle of separation other than that used for the first dimension, and in that each further separation and fractionation is carried out by parallel separating and fractionating methods with the fractions obtained from the preceding separating and fractionating steps.

24. (New) The method according to claim 13, wherein the analysis data for the n-dimensional image of the protein are assembled in a database.--

IN THE ABSTRACT OF THE DISCLOSURE

Cancel the present Abstract of the Disclosure and substitute therefor the enclosed Abstract of the Disclosure which is attached to the substitute specification..

REMARKS

Claims 1-12 have been cancelled and new claims 13-24 have been added. The amendments to the claims have been made only to improve the form of the claims for examination purposes.

The specification and abstract have been amended to conform it to U.S. format.

An early and favorable action on the merits is respectfully requested.

Respectfully submitted,

By: ____________
Jules E. Goldberg
Reg. No. 24,408

January 4, 2002
REED SMITH LLP
375 Park Avenue
New York, NY 10152-1799
JEGforGHK:ram
Enc.: Substitute Specification
Abstract of the Disclosure

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Customer No.	**026418**	
Attorney's Docket No.:	**GK-OEH-119 / 500814.20021**	
U.S. Application No.:		
International Application No.:	**PCT/DE00/02154**	
International Filing Date:	**JULY 4, 2000**	**4 JULY 2000**
Priority Date Claimed:	**JULY 5, 1999**	**5 JULY 1999**
Title of Invention:	**METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME**	
Applicant(s) for (DO/EO/US):	**Thomas MOORE and Anton HORN**	

SUBSTITUTE SPECIFICATION and ABSTRACT

Docket No.: GK-OEH-120/500814.20021

METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority of PCT Application Serial No. PCT/DE00/02154 filed July 4, 2000 and German Application No. 199 32 270.8 filed July 5, 1999, the complete disclosures of which are hereby incorporated by reference.

BACKGROUND OF THE INVENTON

a) Field of the Invention

The invention is directed to a method for the multidimensional analysis of a proteome in which the biological tissue with the proteome to be analyzed is solubilized and the proteins belonging to the proteome are separated, quantitatively determined and identified. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. Special areas of use are opening up in fundamental research, e.g., for clarifying questions pertaining to developmental biology or cell differentiation and in related research for screening active ingredient banks, for the development and optimization of biologically active substances or for differentiating between normal and pathogenic states in organisms.

b) Description of the Related Art

Recently, genomes of organisms have been sequenced completely or in large part [Fraser, C. M. et al.: The minimal gene complement of Mycoplasma genitalium, Science, 1995, Oct. 20, 270 (5235), 397-403; Fleischmann, R. D. et al.: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science, 1995, July 28, 269 (5223), 496-512; Blattner, F. R. et al.: The complete genome sequence of Escherichia coli K-12, Science, 1997, Sept. 5, 277 (5331), 1453-74; Goffeau, A. et al.: Life with 6000 genes, Science, 1996, Oct. 25, 274

(5287), 546, 563-7]. Sequencing of cDNA portions has been even more intensive [Clark, M. S.: Comparative genomics: the key to understanding the Human Genome Project, Bioessays, 1999, Feb. 21 (2), 121-30; Evans, M. J. et al.: Gene trapping and functional genomics, Trends Genet., 1997, Sept. 13 (9), 370-4]. The sequence data are stored in databases. The clarification of the genome of an organism ultimately leads "only" to an understanding of the relatively static information content of the genetic material for this organism. With cDNA sequences, it is possible, in principle, to determine expression levels of the mRNA as they relate to specific cells and specific environments and accordingly to obtain a gene expression pattern of the RNA.

From a gene of the genome, it is possible a) to develop by different processes various mRNA types which code for divergent proteins, and b) to form a large number of extremely differently functioning proteins from this by means of posttranslational modification. Previously known modifications include phosphorylation and dephosphorylation, limited proteolysis, acetylation, methylation, adenylation, sulfation, glycosylation [McDonald, L. J., et al.: Enzymatic and nonenzymatic ADP-ribosylation of cysteins, Mol. Cell. Biochem., 1994 Sept., 138 (1-2), 221-6; Baenziger, J. U.: Protein-specific glycosyltransferases: how and why they do it!, FASEB J., 1994, Oct. 8 (13), 1019-25; Mimnaugh, E. G. et al.: The measurement of ubiquitin and ubiquitinated proteins, Electrophoresis, Feb. 1999, 20 (2), 418-28; Davis, P. J. et al.: Protein modification by thermal processing, Allergy, 1998, 53 (46 Suppl.), 102-5]. However, the expressed and modified proteins ultimately yield the pattern which describes the cell differentiation and the reaction to internal and external influences of cells. Most striking is the limited importance of knowing the genome for the realization of a defined biological state when the various cells in different organs and inside an organ are compared. For example, a liver paranchyma cell, a nerve cell of the brain and a mucosa cell of the intestine have the same set of genetic information but completely different functions brought about by the regulation of the expression of the genome in these cells and the regulation of the enzyme pattern and protein pattern within the cells and the

various tissues.

DNA	RNA	Proteins
Static and descriptive, with exceptions	Transfer of information. Quantity is regulated and transfers the information of the DNA to the protein plane.	Maintaining cell structure, reaction to changes and signals. Interactions with other cells. Quantity and activity are regulated.

The term "proteome" was first used in 1996 [Friedrich, G. A.: Moving beyond the genome projects, Nat. Biotechnol., Oct. 1996, 14 (10), 1234-7].

The proteome, that is, the totality of all proteins in a cell, with a definite development stage and under defined environmental conditions, is a much more dynamic representation of the physiological state of cells, organs and organisms. Proteome analysis investigates which parts of the genome are expressed and modified under defined, cell-specific conditions. This has led to rapidly growing interest in this field, leading to a growing number of publications (PubMed search term: Proteome; over the last 1 year: 64 hits; over the last 2 years: 99 hits; over the last 5 years: 122 hits), conferences and events on this subject.

In order to obtain a quantifiable "picture" of a proteome, the following procedure is currently performed: In a first step, the biological materials must be solubilized and homogenized (exceptions: e.g., in a serum, they are in a homogenous solution). The proteins are isolated or separated in the second step and identified in the third step. In the fourth step, the obtained data are evaluated [Ben, R. H., et al.: Two dimensional electrophoresis, The state of the art and future directions, Proteome Research, New frontiers in functional genomics, Springer 1997, Chap. 2, 13-33].

1. Solubilization

Methods and arrangements known from biochemistry are used for this purpose, e.g., shear homogenizers, ultrasonic processing, high-pressure pressing. The difficulty consists in quantitative solubilization which does not destroy the function of the proteins as far as possible, because only quantitatively solubilized proteins provide a real picture of the specimen material in the subsequent second step (separation and detection of proteins) [Rabilloud, T.: Solubilization of proteins in 2-D electrophoresis, An outline, Methods Mol. Biol;., 1999, 112, 9-19; Rabilloud, T. et al.: Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients, Electrophoresis, Mar.-Apr. 1997, 18 (3-4), 307-16; Staudenmann, W. et al.: Sample Handling for proteome analysis, Electrophoresis, May 1998, 19 (6), 901-8].

2. Separation and detection

At present, two-dimensional gel electrophoresis is essentially used for separating the proteins of the proteome. First tests with two-dimensional HPLC have been carried out. However, they have not yet achieved the separation effect of two-dimensional electrophoresis [Opiteck G. J. et al.: Comprehensive two-dimensional high-performance liquid chromatography for the separation of overexpressed proteins and proteome mapping, Anal. Biochem. May 1998, 1; 258 (2): 349-61]. The first dimension of two-dimensional electrophoresis is isolation according to the isoelectric point, that is, ultimately, according to the charge characteristics of a protein. In the second dimension, the proteins are separated according to size in a denaturing sodium dodecyl sulfate gel. This separation technique has been known for about 20 years. An advantage of two-dimensional electrophoresis consists in the possibility of separating a relatively large number of proteins on a surface with high resolution. Presently, it is assumed that approximately 10,000 proteins can be detected in a two-dimensional gel of this kind [Klose, J. et al.: Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome, Electrophoresis, 1995, June

16 (6), 1034-59]. Another advantage is that it is possible to quantify the separated proteins by radioactive marking or after staining with techniques that are likewise known. These quantification methods are protein-specific, have a limited dynamic detection range, are generally difficult to automate and are dependent on the respective conditions of use (which often can not be reproduced) [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]. They are only suitable for relative determinations. Quantification by immunological characteristics is problematic because blot techniques having limited meaningfulness in terms of quantitative information must be used for this purpose.

This results in a fingerprint-like pattern which characterizes the proteome.

This separation technique has the following disadvantages:

- limited dynamic range due to the load capacity of the separating gel
- the maximum quantity of proteins that may be used is limited to a range of µg to mg protein [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]
- restriction of sample volume used
- separation is limited to two dimensions
- the ampholytes required for separation and the acrylamide gel material can lead to artifacts and can accordingly contribute to misinterpretations which are difficult to detect
- proteins that are present in very high concentrations result in relatively strong signals and overlap proteins in low concentrations, so that direct identification and quantification is impossible in this case
- the loss of the native conformation in denaturing separating gel causes the loss of biologically functional characteristics and impedes the identification of proteins by determining their biological characteristics, for example, their catalytic activity or specific bonding characteristics
- secondary analysis, such as the frequently used specific proteolysis of individual proteins, followed by determinations of mass necessitates a step for

extracting from the gel or blot membrane which is difficult to automate.

3. Identification of proteins

Sequencing, mass analysis and estimation of the isoelectric point from the length of run in the gel and mass analysis of peptide fragments after separation from the gel and tryptic digestion in mass spectrometry are normally used for this purpose [Shevchenko, A. et al.: Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two-dimensional gels, Proc. Natl. Acad. Sci. USA, 1996, Dec. 10, 93 (25), 14440-5; Traini, M. et al.: Towards an automated approach for protein identification in proteome projects, Electrophoresis, 1998, Aug. 19 (11), 1941-9]. Features such as the catalytic activity of the proteins and the native conformation are almost completely excluded from the utilized separating technique and are not available for identification.

In particular, the known identification methods have the following advantages and disadvantages:

- The sequencing is carried out by Edman degradation in automated arrangements and is relatively costly and time-consuming. It requires greater quantities of the protein. Therefore, in spite of current further development for mass screening, it is less suitable [Gooley, A. A. et al.: A role for Edman degradation in proteome studies, Electrophoresis, 1997, June 18(7), 1068-72]. However, this analytic step is necessary in most cases for identification of primarily unknown proteins.

- The specificity of information of mass determination of a protein which should finally lead to its identification is increased in that the proteins undergo protease digestion after separation and the information obtained by means of mass analysis is compared with the masses of the peptide sequences predicted from the primary structure after tryptic digestion. Essentially two types of mass spectrometry are used: The first is Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) and the second is ElectroSpray Ionization Mass Spectrometry (ESI-MS) [Ducret, A. et al.: High Throughput protein characterization

by automated reverse-phase chromatography/electrospray tandem mass spectrometry, Protein Sci., 1998, Mar 7 (3), 706-19; Parker, K. C. et al.: Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination, Electrophoresis, 1998, Aug. 19 (11), 1920-32]. The first method has the advantage that it allows a very large mass range of up to 1 million Dalton to be analyzed and can be carried out in a relatively robust manner. However, it can be carried out only discontinuously. The ESI technique, on the other hand, can be appended almost continuously to separating techniques and is presently showing a sharp growth in the development of breadth of application as well as technological possibilities. The enormous advances achieved in recent years with both techniques allow mass resolutions to isotope distribution, that is, resolutions of less than 1 Dalton. In this way, a mass spectrum of peptide fragments is obtained according to sequence-specific, defined protease digestion or another defined splitting of the proteins. This spectrum is typical for every protein and is used for protein identification in sequence databases of proteins and expressed sequence tag banks. Since the identification of the protein by precise identification of the predicted peptides takes place after protease digestion, any posttranslational modification of the proteins, e.g., by glycosylation, interferes with detection. Further, fragmentation spectra of the individual peptides in the mass spectrometer can supply information about the amino acid sequence of the peptides. This sequence information can be used by itself or along with the other known protein data to identify this protein in a sequence database. This method of sequence analysis is not yet routinely used at present due to the difficulties of correct data interpretation. The limits of protein identification through mass spectrometry methods reside in the incomplete detection of all protein sequences in existing databases.

4. Data analysis

The characteristics of the individual detected proteins from separation in two-dimensional electrophoresis, such as quantity, isoelectric point and size, and the data for protein identification from additional steps, e.g., sequencing or mass

spectrometry, are combined. This produces the picture of the totality of the proteins with their identity and quantity in the respective proteome.

OBJECT AND SUMMARY OF THE INVENTION

It is the primary object of the invention to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time.

According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the m_1 liquid fractions obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by τ different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that after combining the analysis data an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data space.

The method according to the invention is not subject to the tight limitation on quantity due to the load capacity of previously used two-dimensional electrophoresis. Protein quantities in the range of several grams can be used. The separating matrices can be utilized repeatedly. In this way, greater reproducibility of results can be achieved. The sample material that is used is in liquid phase and is accordingly immediately accessible for subsequent analysis steps. The improved maintaining of native characteristics during separation makes possible analytic methods such as activity determination and immunological processes based on the native conformation of the analytes. The separation of analytes with the same charge characteristics and size characteristics is not possible in the two-dimensional electrophoresis that is usually used. However, this restriction is eliminated through the use of at least one further characteristic, such as the hydrophobicity of the analytes, for separation. After separation, the samples in fractions are also available

for additional preparative tasks.

The invention will be described more fully in the following with reference to an embodiment example shown in the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings:

Fig. 1 shows the separation of 1000 proteins in three dimensions. Figure 1 comprises:

Fig. 1a: fractions 1 to 33

Fig. 1b: fractions 33/34 to 67

Fig. 1c: fractions 68 to 100;

Fig. 2 shows a graphic three-dimensional view of the fractions according to Fig. 1.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

As an embodiment example, 1000 proteins are to be described by three characteristics A, B, C. These characteristics may be, e.g., size, charge and hydrophobicity. The characteristics are randomly distributed in the proteins. All proteins are numbered consecutively. Subsequently, separation is carried out according to characteristic A (e.g., size), resulting in 100 fractions a with the corresponding proteins. These fractions a are separated into 10 fractions b according to characteristic B (e.g., charge).

Each of these fractions b is subjected to separation based on characteristic C (e.g., hydrophobicity) and gives fractions c. In total, 100 x 10 x 10 = 10,000 individual fractions are obtained. Each protein obtained by separation is uniquely allocated to one of the fractions a, b, c according to its characteristics. In the assignment according to Fig. 1, the respective fractions are designated by number. In this case, the fractions a are associated with characteristic A. They divide the possible value range of characteristic A into one hundred equal parts, i.e., assuming a value range from 0 to 100, value 1, for example, corresponds to range 0

to 1, value 2 corresponds to range 1 to 2, ..., and value 100 corresponds to range 99 to 100. Analogously, the possible value ranges of characteristics B and C are divided into ten equal parts, i.e., value 1, for example, corresponds to range 1 to 10. On the average, every tenth fraction contains a protein.

Considered at random, there is a possibility of multiple assignments. In the example shown in the list according to Figs. 1a-c, there are 39 double occupancies and one triple occupancy of fractions.

For reasons of space and for the sake of clarity, the empty 9,000 fractions are not shown.

Fig. 1 contains the following list in tabular form:

Protein No.	Fractions a	Fractions b	Fractions c

Fig. 1a shows fractions a = 1 to 33, Fig. 2a shows fractions a = 33/34 to 67 and Fig. 1c shows fractions a = 68 to 100. Fig. 2 shows a three-dimensional diagram with the positions of the fractions occupied by proteins according to Fig. 1

While the foregoing description and drawings represent the present invention, it will be obvious to those skilled in the art that various changes may be made therein without departing from the true spirit and scope of the present invention.

Assignment of Reference Numbers

A, B, C	-	characteristic of proteins
a, b, c	-	fraction

ABSTRACT OF THE DISCLOSURE

The invention is directed to a method for the multidimensional analysis of a proteome. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. The object of the invention is to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time. According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the liquid fractions m_1 obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by o different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that, after combining the analysis data in a database, an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data network.

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Customer No.	**026418**	
Attorney's Docket No.:	**GK-OEH-119 / 500814.20021**	
U.S. Application No.:		
International Application No.:	**PCT/DE00/02154**	
International Filing Date:	**JULY 4, 2000**	**4 JULY 2000**
Priority Date Claimed:	**JULY 5, 1999**	**5 JULY 1999**
Title of Invention:	**METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME**	
Applicant(s) for (DO/EO/US):	**Thomas MOORE and Anton HORN**	

MARKED-UP/BOLDED

SUBSTITUTE

SPECIFICATION

and

ABSTRACT

Docket No.: GK-OEH-120/500814.20021

METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME

CROSS-REFERENCE TO RELATED APPLICATIONS

This application claims priority of PCT Application Serial No. PCT/DE00/02154 filed July 4, 2000 and German Application No. 199 32 270.8 filed July 5, 1999, the complete disclosures of which are hereby incorporated by reference.

BACKGROUND OF THE INVENTON

a) Field of the Invention

The invention is directed to a method for the multidimensional analysis of a proteome in which the biological tissue with the proteome to be analyzed is solubilized and the proteins belonging to the proteome are separated, quantitatively determined and identified. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. Special areas of use are opening up in fundamental research, e.g., for clarifying questions pertaining to developmental biology or cell differentiation and in related research for screening active ingredient banks, for the development and optimization of biologically active substances or for differentiating between normal and pathogenic states in organisms.

b) Description of the Related Art

Recently, genomes of organisms have been sequenced completely or in large part [Fraser, C. M. et al.: The minimal gene complement of Mycoplasma genitalium, Science, 1995, Oct. 20, 270 (5235), 397-403; Fleischmann, R. D. et al.: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science, 1995, July 28, 269 (5223), 496-512; Blattner, F. R. et al.: The complete genome sequence of Escherichia coli K-12, Science, 1997, Sept. 5, 277 (5331),

1453-74; Goffeau, A. et al.: Life with 6000 genes, Science, 1996, Oct. 25, 274 (5287), 546, 563-7]. Sequencing of cDNA portions has been even more intensive [Clark, M. S.: Comparative genomics: the key to understanding the Human Genome Project, Bioessays, 1999, Feb. 21 (2), 121-30; Evans, M. J. et al.: Gene trapping and functional genomics, Trends Genet., 1997, Sept. 13 (9), 370-4]. The sequence data are stored in databases. The clarification of the genome of an organism ultimately leads "only" to an understanding of the relatively static information content of the genetic material for this organism. With cDNA sequences, it is possible, in principle, to determine expression levels of the mRNA as they relate to specific cells and specific environments and accordingly to obtain a gene expression pattern of the RNA.

From a gene of the genome, it is possible a) to develop by different processes various mRNA types which code for divergent proteins, and b) to form a large number of extremely differently functioning proteins from this by means of posttranslational modification. Previously known modifications include phosphorylation and dephosphorylation, limited proteolysis, acetylation, methylation, adenylation, sulfation, glycosylation [McDonald, L. J., et al.: Enzymatic and nonenzymatic ADP-ribosylation of cysteins, Mol. Cell. Biochem., 1994 Sept., 138 (1-2), 221-6; Baenziger, J. U.: Protein-specific glycosyltransferases: how and why they do it!, FASEB J., 1994, Oct. 8 (13), 1019-25; Mimnaugh, E. G. et al.: The measurement of ubiquitin and ubiquitinated proteins, Electrophoresis, Feb. 1999, 20 (2), 418-28; Davis, P. J. et al.: Protein modification by thermal processing, Allergy, 1998, 53 (46 Suppl.), 102-5]. However, the expressed and modified proteins ultimately yield the pattern which describes the cell differentiation and the reaction to internal and external influences of cells. Most striking is the limited importance of knowing the genome for the realization of a defined biological state when the various cells in different organs and inside an organ are compared. For example, a liver paranchyma cell, a nerve cell of the brain and a mucosa cell of the intestine have the same set of genetic information but completely different functions

brought about by the regulation of the expression of the genome in these cells and the regulation of the enzyme pattern and protein pattern within the cells and the various tissues.

DNA	RNA	Proteins
Static and descriptive, with exceptions	Transfer of information. Quantity is regulated and transfers the information of the DNA to the protein plane.	Maintaining cell structure, reaction to changes and signals. Interactions with other cells. Quantity and activity are regulated.

The term "proteome" was first used in 1996 [Friedrich, G. A.: Moving beyond the genome projects, Nat. Biotechnol., Oct. 1996, 14 (10), 1234-7].

The proteome, that is, the totality of all proteins in a cell, with a definite development stage and under defined environmental conditions, is a much more dynamic representation of the physiological state of cells, organs and organisms. Proteome analysis investigates which parts of the genome are expressed and modified under defined, cell-specific conditions. This has led to rapidly growing interest in this field, leading to a growing number of publications (PubMed search term: Proteome; over the last 1 year: 64 hits; over the last 2 years: 99 hits; over the last 5 years: 122 hits), conferences and events on this subject.

In order to obtain a quantifiable "picture" of a proteome, the following procedure is currently performed: In a first step, the biological materials must be solubilized and homogenized (exceptions: e.g., in a serum, they are in a homogenous solution). The proteins are isolated or separated in the second step and identified in the third step. In the fourth step, the obtained data are evaluated [Ben, R. H., et al.: Two dimensional electrophoresis, The state of the art and future

directions, Proteome Research, New frontiers in functional genomics, Springer 1997, Chap. 2, 13-33].

1. Solubilization

Methods and arrangements known from biochemistry are used for this purpose, e.g., shear homogenizers, ultrasonic processing, high-pressure pressing. The difficulty consists in quantitative solubilization which does not destroy the function of the proteins as far as possible, because only quantitatively solubilized proteins provide a real picture of the specimen material in the subsequent second step (separation and detection of proteins) [Rabilloud, T.: Solubilization of proteins in 2-D electrophoresis, An outline, Methods Mol. Biol;., 1999, 112, 9-19; Rabilloud, T. et al.: Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients, Electrophoresis, Mar.-Apr. 1997, 18 (3-4), 307-16; Staudenmann, W. et al.: Sample Handling for proteome analysis, Electrophoresis, May 1998, 19 (6), 901-8].

2. Separation and detection

At present, two-dimensional gel electrophoresis is essentially used for separating the proteins of the proteome. First tests with two-dimensional HPLC have been carried out. However, they have not yet achieved the separation effect of two-dimensional electrophoresis [Opiteck G. J. et al.: Comprehensive two-dimensional high-performance liquid chromatography for the separation of overexpressed proteins and proteome mapping, Anal. Biochem. May 1998, 1; 258 (2): 349-61]. The first dimension of two-dimensional electrophoresis is isolation according to the isoelectric point, that is, ultimately, according to the charge characteristics of a protein. In the second dimension, the proteins are separated according to size in a denaturing sodium dodecyl sulfate gel. This separation technique has been known for about 20 years. An advantage of two-dimensional electrophoresis consists in the possibility of separating a relatively large number of

proteins on a surface with high resolution. Presently, it is assumed that approximately 10,000 proteins can be detected in a two-dimensional gel of this kind [Klose, J. et al.: Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome, Electrophoresis, 1995, June 16 (6), 1034-59]. Another advantage is that it is possible to quantify the separated proteins by radioactive marking or after staining with techniques that are likewise known. These quantification methods are protein-specific, have a limited dynamic detection range, are generally difficult to automate and are dependent on the respective conditions of use (which often can not be reproduced) [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]. They are only suitable for relative determinations. Quantification by immunological characteristics is problematic because blot techniques having limited meaningfulness in terms of quantitative information must be used for this purpose.

This results in a fingerprint-like pattern which characterizes the proteome.

This separation technique has the following disadvantages:

- limited dynamic range due to the load capacity of the separating gel
- the maximum quantity of proteins that may be used is limited to a range of µg to mg protein [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]
- restriction of sample volume used
- separation is limited to two dimensions
- the ampholytes required for separation and the acrylamide gel material can lead to artifacts and can accordingly contribute to misinterpretations which are difficult to detect
- proteins that are present in very high concentrations result in relatively strong signals and overlap proteins in low concentrations, so that direct identification and quantification is impossible in this case
- the loss of the native conformation in denaturing separating gel causes the

loss of biologically functional characteristics and impedes the identification of proteins by determining their biological characteristics, for example, their catalytic activity or specific bonding characteristics

- secondary analysis, such as the frequently used specific proteolysis of individual proteins, followed by determinations of mass necessitates a step for extracting from the gel or blot membrane which is difficult to automate.

3. Identification of proteins

Sequencing, mass analysis and estimation of the isoelectric point from the length of run in the gel and mass analysis of peptide fragments after separation from the gel and tryptic digestion in mass spectrometry are normally used for this purpose [Shevchenko, A. et al.: Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two-dimensional gels, Proc. Natl. Acad. Sci. USA, 1996, Dec. 10, 93 (25), 14440-5; Traini, M. et al.: Towards an automated approach for protein identification in proteome projects, Electrophoresis, 1998, Aug. 19 (11), 1941-9]. Features such as the catalytic activity of the proteins and the native conformation are almost completely excluded from the utilized separating technique and are not available for identification.

In particular, the known identification methods have the following advantages and disadvantages:

- The sequencing is carried out by Edman degradation in automated arrangements and is relatively costly and time-consuming. It requires greater quantities of the protein. Therefore, in spite of current further development for mass screening, it is less suitable [Gooley, A. A. et al.: A role for Edman degradation in proteome studies, Electrophoresis, 1997, June 18(7), 1068-72]. However, this analytic step is necessary in most cases for identification of primarily unknown proteins.
- The specificity of information of mass determination of a protein which should finally lead to its identification is increased in that the proteins undergo

protease digestion after separation and the information obtained by means of mass analysis is compared with the masses of the peptide sequences predicted from the primary structure after tryptic digestion. Essentially two types of mass spectrometry are used: The first is Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) and the second is ElectroSpray Ionization Mass Spectrometry (ESI-MS) [Ducret, A. et al.: High Throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry, Protein Sci., 1998, Mar 7 (3), 706-19; Parker, K. C. et al.: Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination, Electrophoresis, 1998, Aug. 19 (11), 1920-32]. The first method has the advantage that it allows a very large mass range of up to 1 million Dalton to be analyzed and can be carried out in a relatively robust manner. However, it can be carried out only discontinuously. The ESI technique, on the other hand, can be appended almost continuously to separating techniques and is presently showing a sharp growth in the development of breadth of application as well as technological possibilities. The enormous advances achieved in recent years with both techniques allow mass resolutions to isotope distribution, that is, resolutions of less than 1 Dalton. In this way, a mass spectrum of peptide fragments is obtained according to sequence-specific, defined protease digestion or another defined splitting of the proteins. This spectrum is typical for every protein and is used for protein identification in sequence databases of proteins and expressed sequence tag banks. Since the identification of the protein by precise identification of the predicted peptides takes place after protease digestion, any posttranslational modification of the proteins, e.g., by glycosylation, interferes with detection. Further, fragmentation spectra of the individual peptides in the mass spectrometer can supply information about the amino acid sequence of the peptides. This sequence information can be used by itself or along with the other known protein data to identify this protein in a sequence database. This method of sequence analysis is not yet routinely used at present due to the difficulties of correct data interpretation. The limits of protein

identification through mass spectrometry methods reside in the incomplete detection of all protein sequences in existing databases.

4. Data analysis

The characteristics of the individual detected proteins from separation in two-dimensional electrophoresis, such as quantity, isoelectric point and size, and the data for protein identification from additional steps, e.g., sequencing or mass spectrometry, are combined. This produces the picture of the totality of the proteins with their identity and quantity in the respective proteome.

OBJECT AND SUMMARY OF THE INVENTION

It is the **primary** object of the invention to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time.

According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the m_1 liquid fractions obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by τ different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that after combining the analysis data an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data space.

[Advantageous embodiment forms of the method are set forth in the subclaims 2 to 12.]

The method according to the invention is not subject to the tight limitation on quantity due to the load capacity of previously used two-dimensional electrophoresis. Protein quantities in the range of several grams can be used. The

separating matrices can be utilized repeatedly. In this way, greater reproducibility of results can be achieved. The sample material that is used is in liquid phase and is accordingly immediately accessible for subsequent analysis steps. The improved maintaining of native characteristics during separation makes possible analytic methods such as activity determination and immunological processes based on the native conformation of the analytes. The separation of analytes with the same charge characteristics and size characteristics is not possible in the two-dimensional electrophoresis that is usually used. However, this restriction is eliminated through the use of at least one further characteristic, such as the hydrophobicity of the analytes, for separation. After separation, the samples in fractions are also available for additional preparative tasks.

The invention will be described more fully in the following with reference to an embodiment example shown in the drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

In the drawings:

Fig. 1 shows the separation of 1000 proteins in three dimensions.

Figure 1 comprises:

Fig. 1a: fractions 1 to 33

Fig. **1b [2a]**: fractions 33/34 to 67

Fig. **1c [3a]**: fractions 68 to 100;

Fig. 2 shows a graphic three-dimensional view of the fractions according to Fig. 1.

DESCRIPTION OF THE PREFERRED EMBODIMENTS

As an embodiment example, 1000 proteins are to be described by three characteristics A, B, C. These characteristics may be, e.g., size, charge and hydrophobicity. The characteristics are randomly distributed in the proteins. All proteins are numbered consecutively. Subsequently, separation is carried out

according to characteristic A (e.g., size), resulting in 100 fractions a with the corresponding proteins. These fractions a are separated into 10 fractions b according to characteristic B (e.g., charge).

Each of these fractions b is subjected to separation based on characteristic C (e.g., hydrophobicity) and gives fractions c. In total, 100 x 10 x 10 = 10,000 individual fractions are obtained. Each protein obtained by separation is uniquely allocated to one of the fractions a, b, c according to its characteristics. In the assignment according to Fig. 1, the respective fractions are designated by number. In this case, the fractions a are associated with characteristic A. They divide the possible value range of characteristic A into one hundred equal parts, i.e., assuming a value range from 0 to 100, value 1, for example, corresponds to range 0 to 1, value 2 corresponds to range 1 to 2, ..., and value 100 corresponds to range 99 to 100. Analogously, the possible value ranges of characteristics B and C are divided into ten equal parts, i.e., value 1, for example, corresponds to range 1 to 10. On the average, every tenth fraction contains a protein.

Considered at random, there is a possibility of multiple assignments. In the example shown in the list according to Figs. 1a-c, there are 39 double occupancies and one triple occupancy of fractions.

For reasons of space and for the sake of clarity, the empty 9,000 fractions are not shown.

Fig. 1 contains the following list in tabular form:

Protein No.	Fractions a	Fractions b	Fractions c

Fig. 1a shows fractions a = 1 to 33, Fig. 2a shows fractions a = 33/34 to 67 and Fig. 1c shows fractions a = 68 to 100. Fig. 2 shows a three-dimensional diagram with the positions of the fractions occupied by proteins according to Fig. 1

While the foregoing description and drawings represent the

present invention, it will be obvious to those skilled in the art that various changes may be made therein without departing from the true spirit and scope of the present invention.

Abstract **of the Disclosure**

The invention is directed to a method for the multidimensional analysis of a proteome. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. It is the object of the invention to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time. According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the liquid fractions m_1 obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by o different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that, after combining the analysis data in a database, an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data network.

MARKED-UP/BOLDED SUBSTITUTE SPECIFICATION

Description of the Invention

METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME

The invention is directed to a method for the multidimensional analysis of a proteome in which the biological tissue with the proteome to be analyzed is solubilized and the proteins belonging to the proteome are separated, quantitatively determined and identified. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. Special areas of use are opening up in fundamental research, e.g., for clarifying questions pertaining to developmental biology or cell differentiation and in related research for screening active ingredient banks, for the development and optimization of biologically active substances or for differentiating between normal and pathogenic states in organisms.

Recently, genomes of organisms have been sequenced completely or in large part [Fraser, C. M. et al.: The minimal gene complement of Mycoplasma genitalium, Science, 1995, Oct. 20, 270 (5235), 397-403; Fleischmann, R. D. et al.: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science, 1995, July 28, 269 (5223), 496-512; Blattner, F. R. et al.: The complete genome sequence of Escherichia coli K-12, Science, 1997, Sept. 5, 277 (5331), 1453-74; Goffeau, A. et al.: Life with 6000 genes, Science, 1996, Oct. 25, 274 (5287), 546, 563-7]. Sequencing of cDNA portions has been even more intensive [Clark, M. S.: Comparative genomics: the key to understanding the Human Genome Project, Bioessays, 1999, Feb. 21 (2), 121-30; Evans, M. J. et al.: Gene trapping and functional genomics, Trends Genet., 1997, Sept. 13 (9), 370-4]. The sequence data are stored in databases. The clarification of the genome of an organism ultimately leads "only" to an understanding of the relatively static information content of the genetic material for this organism. With cDNA sequences, it is possible, in principle, to determine expression levels of the mRNA as they relate to specific cells and specific environments and accordingly to obtain a gene expression pattern of the RNA.

From a gene of the genome, it is possible a) to develop by different processes various mRNA types which code for divergent proteins, and b) to form a

large number of extremely differently functioning proteins from this by means of posttranslational modification. Previously known modifications include phosphorylation and dephosphorylation, limited proteolysis, acetylation, methylation, adenylation, sulfation, glycosylation [McDonald, L. J., et al.: Enzymatic and nonenzymatic ADP-ribosylation of cysteins, Mol. Cell. Biochem., 1994 Sept., 138 (1-2), 221-6; Baenziger, J. U.: Protein-specific glycosyltransferases: how and why they do it!, FASEB J., 1994, Oct. 8 (13), 1019-25; Mimnaugh, E. G. et al.: The measurement of ubiquitin and ubiquitinated proteins, Electrophoresis, Feb. 1999, 20 (2), 418-28; Davis, P. J. et al.: Protein modification by thermal processing, Allergy, 1998, 53 (46 Suppl.), 102-5]. However, the expressed and modified proteins ultimately yield the pattern which describes the cell differentiation and the reaction to internal and external influences of cells. Most striking is the limited importance of knowing the genome for the realization of a defined biological state when the various cells in different organs and inside an organ are compared. For example, a liver paranchyma cell, a nerve cell of the brain and a mucosa cell of the intestine have the same set of genetic information but completely different functions brought about by the regulation of the expression of the genome in these cells and the regulation of the enzyme pattern and protein pattern within the cells and the various tissues.

DNA	RNA	Proteins
Static and descriptive, with exceptions	Transfer of information. Quantity is regulated and transfers the information of the DNA to the protein plane.	Maintaining cell structure, reaction to changes and signals. Interactions with other cells. Quantity and activity are regulated.

The term "proteome" was first used in 1996 [Friedrich, G. A.: Moving beyond the genome projects, Nat. Biotechnol., Oct. 1996, 14 (10), 1234-7]. The proteome, that is, the totality of all proteins in a cell, with a

definite development stage and under defined environmental conditions, is a much more dynamic representation of the physiological state of cells, organs and organisms. Proteome analysis investigates which parts of the genome are expressed and modified under defined, cell-specific conditions. This has led to rapidly growing interest in this field, leading to a growing number of publications (PubMed search term: Proteome; over the last 1 year: 64 hits; over the last 2 years: 99 hits; over the last 5 years: 122 hits), conferences and events on this subject.

In order to obtain a quantifiable "picture" of a proteome, the following procedure is currently performed: In a first step, the biological materials must be solubilized and homogenized (exceptions: e.g., in a serum, they are in a homogenous solution). The proteins are isolated or separated in the second step and identified in the third step. In the fourth step, the obtained data are evaluated [Ben, R. H., et al.: Two dimensional electrophoresis, The state of the art and future directions, Proteome Research, New frontiers in functional genomics, Springer 1997, Chap. 2, 13-33].

1. Solubilization

Methods and arrangements known from biochemistry are used for this purpose, e.g., shear homogenizers, ultrasonic processing, high-pressure pressing. The difficulty consists in quantitative solubilization which does not destroy the function of the proteins as far as possible, because only quantitatively solubilized proteins provide a real picture of the specimen material in the subsequent second step (separation and detection of proteins) [Rabilloud, T.: Solubilization of proteins in 2-D electrophoresis, An outline, Methods Mol. Biol;., 1999, 112, 9-19; Rabilloud, T. et al.: Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients, Electrophoresis, Mar.-Apr. 1997, 18 (3-4), 307-16; Staudenmann, W. et al.: Sample Handling for proteome analysis, Electrophoresis, May 1998, 19 (6), 901-8].

2. Separation and detection

At present, two-dimensional gel electrophoresis is essentially used for separating the proteins of the proteome. First tests with two-dimensional HPLC

have been carried out. However, they have not yet achieved the separation effect of two-dimensional electrophoresis [Opiteck G. J. et al.: Comprehensive two-dimensional high-performance liquid chromatography for the separation of overexpressed proteins and proteome mapping, Anal. Biochem. May 1998, 1; 258 (2): 349-61]. The first dimension of two-dimensional electrophoresis is isolation according to the isoelectric point, that is, ultimately, according to the charge characteristics of a protein. In the second dimension, the proteins are separated according to size in a denaturing sodium dodecyl sulfate gel. This separation technique has been known for about 20 years. An advantage of two-dimensional electrophoresis consists in the possibility of separating a relatively large number of proteins on a surface with high resolution. Presently, it is assumed that approximately 10,000 proteins can be detected in a two-dimensional gel of this kind [Klose, J. et al.: Two-dimensional electrophoresis of proteins: an updated protocol and implications for a functional analysis of the genome, Electrophoresis, 1995, June 16 (6), 1034-59]. Another advantage is that it is possible to quantify the separated proteins by radioactive marking or after staining with techniques that are likewise known. These quantification methods are protein-specific, have a limited dynamic detection range, are generally difficult to automate and are dependent on the respective conditions of use (which often can not be reproduced) [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]. They are only suitable for relative determinations. Quantification by immunological characteristics is problematic because blot techniques having limited meaningfulness in terms of quantitative information must be used for this purpose.

This results in a fingerprint-like pattern which characterizes the proteome.

This separation technique has the following disadvantages:

- limited dynamic range due to the load capacity of the separating gel
- the maximum quantity of proteins that may be used is limited to a range of µg to mg protein [James, P.: Of genomes and proteomes, Biochem. Biophys. Res. Commun., 1997, Feb. 3, 231 (1), 1-6]
- restriction of sample volume used
- separation is limited to two dimensions

- the ampholytes required for separation and the acrylamide gel material can lead to artifacts and can accordingly contribute to misinterpretations which are difficult to detect
- proteins that are present in very high concentrations result in relatively strong signals and overlap proteins in low concentrations, so that direct identification and quantification is impossible in this case
- the loss of the native conformation in denaturing separating gel causes the loss of biologically functional characteristics and impedes the identification of proteins by determining their biological characteristics, for example, their catalytic activity or specific bonding characteristics
- secondary analysis, such as the frequently used specific proteolysis of individual proteins, followed by determinations of mass necessitates a step for extracting from the gel or blot membrane which is difficult to automate.

3. Identification of proteins

Sequencing, mass analysis and estimation of the isoelectric point from the length of run in the gel and mass analysis of peptide fragments after separation from the gel and tryptic digestion in mass spectrometry are normally used for this purpose [Shevchenko, A. et al.: Linking genome and proteome by mass spectrometry: large-scale identification of yeast proteins from two-dimensional gels, Proc. Natl. Acad. Sci. USA, 1996, Dec. 10, 93 (25), 14440-5; Traini, M. et al.: Towards an automated approach for protein identification in proteome projects, Electrophoresis, 1998, Aug. 19 (11), 1941-9]. Features such as the catalytic activity of the proteins and the native conformation are almost completely excluded from the utilized separating technique and are not available for identification.

In particular, the known identification methods have the following advantages and disadvantages:

- The sequencing is carried out by Edman degradation in automated arrangements and is relatively costly and time-consuming. It requires greater quantities of the protein. Therefore, in spite of current further development for mass screening, it is less suitable [Gooley, A. A. et al.: A role for Edman degradation in proteome studies, Electrophoresis, 1997, June 18(7), 1068-72]. However, this

analytic step is necessary in most cases for identification of primarily unknown proteins.

- The specificity of information of mass determination of a protein which should finally lead to its identification is increased in that the proteins undergo protease digestion after separation and the information obtained by means of mass analysis is compared with the masses of the peptide sequences predicted from the primary structure after tryptic digestion. Essentially two types of mass spectrometry are used: The first is Matrix Assisted Laser Desorption Ionization Mass Spectrometry (MALDI-MS) and the second is ElectroSpray Ionization Mass Spectrometry (ESI-MS) [Ducret, A. et al.: High Throughput protein characterization by automated reverse-phase chromatography/electrospray tandem mass spectrometry, Protein Sci., 1998, Mar 7 (3), 706-19; Parker, K. C. et al.: Identification of yeast proteins from two-dimensional gels: working out spot cross-contamination, Electrophoresis, 1998, Aug. 19 (11), 1920-32]. The first method has the advantage that it allows a very large mass range of up to 1 million Dalton to be analyzed and can be carried out in a relatively robust manner. However, it can be carried out only discontinuously. The ESI technique, on the other hand, can be appended almost continuously to separating techniques and is presently showing a sharp growth in the development of breadth of application as well as technological possibilities. The enormous advances achieved in recent years with both techniques allow mass resolutions to isotope distribution, that is, resolutions of less than 1 Dalton. In this way, a mass spectrum of peptide fragments is obtained according to sequence-specific, defined protease digestion or another defined splitting of the proteins. This spectrum is typical for every protein and is used for protein identification in sequence databases of proteins and expressed sequence tag banks. Since the identification of the protein by precise identification of the predicted peptides takes place after protease digestion, any posttranslational modification of the proteins, e.g., by glycosylation, interferes with detection. Further, fragmentation spectra of the individual peptides in the mass spectrometer can supply information about the amino acid sequence of the peptides. This sequence information can be used by itself or along with the other known protein data to identify this protein in a sequence database. This method of sequence analysis is not yet routinely used at

present due to the difficulties of correct data interpretation. The limits of protein identification through mass spectrometry methods reside in the incomplete detection of all protein sequences in existing databases.

4. Data analysis

The characteristics of the individual detected proteins from separation in two-dimensional electrophoresis, such as quantity, isoelectric point and size, and the data for protein identification from additional steps, e.g., sequencing or mass spectrometry, are combined. This produces the picture of the totality of the proteins with their identity and quantity in the respective proteome.

It is the object of the invention to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time.

According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the m_1 liquid fractions obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by τ different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that after combining the analysis data an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data space.

Advantageous embodiment forms of the method are set forth in the subclaims 2 to 12.

The method according to the invention is not subject to the tight limitation on quantity due to the load capacity of previously used two-dimensional electrophoresis. Protein quantities in the range of several grams can be used. The separating matrices can be utilized repeatedly. In this way, greater reproducibility of results can be achieved. The sample material that is used is in liquid phase and is accordingly immediately accessible for subsequent analysis steps. The improved maintaining of native characteristics during separation makes possible analytic

methods such as activity determination and immunological processes based on the native conformation of the analytes. The separation of analytes with the same charge characteristics and size characteristics is not possible in the two-dimensional electrophoresis that is usually used. However, this restriction is eliminated through the use of at least one further characteristic, such as the hydrophobicity of the analytes, for separation. After separation, the samples in fractions are also available for additional preparative tasks.

The invention will be described more fully in the following with reference to an embodiment example shown in the drawing.

Fig. 1 shows the separation of 1000 proteins in three dimensions

Fig. 1a: fractions 1 to 33

Fig. 2a: fractions 33/34 to 67

Fig. 3a: fractions 68 to 100;

Fig. 2 shows a graphic three-dimensional view of the fractions according to Fig. 1.

As an embodiment example, 1000 proteins are to be described by three characteristics A, B, C. These characteristics may be, e.g., size, charge and hydrophobicity. The characteristics are randomly distributed in the proteins. All proteins are numbered consecutively. Subsequently, separation is carried out according to characteristic A (e.g., size), resulting in 100 fractions a with the corresponding proteins. These fractions a are separated into 10 fractions b according to characteristic B (e.g., charge).

Each of these fractions b is subjected to separation based on characteristic C (e.g., hydrophobicity) and gives fractions c. In total, 100 x 10 x 10 = 10,000 individual fractions are obtained. Each protein obtained by separation is uniquely allocated to one of the fractions a, b, c according to its characteristics. In the assignment according to Fig. 1, the respective fractions are designated by number. In this case, the fractions a are associated with characteristic A. They divide the possible value range of characteristic A into one hundred equal parts, i.e., assuming a value range from 0 to 100, value 1, for example, corresponds to range 0

to 1, value 2 corresponds to range 1 to 2, ..., and value 100 corresponds to range 99 to 100. Analogously, the possible value ranges of characteristics B and C are divided into ten equal parts, i.e., value 1, for example, corresponds to range 1 to 10. On the average, every tenth fraction contains a protein.

Considered at random, there is a possibility of multiple assignments. In the example shown in the list according to Figs. 1a-c, there are 39 double occupancies and one triple occupancy of fractions.

For reasons of space and for the sake of clarity, the empty 9,000 fractions are not shown.

Fig. 1 contains the following list in tabular form:

Protein No.	Fractions a	Fractions b	Fractions c

Fig. 1a shows fractions a = 1 to 33, Fig. 2a shows fractions a = 33/34 to 67 and Fig. 1c shows fractions a = 68 to 100. Fig. 2 shows a three-dimensional diagram with the positions of the fractions occupied by proteins according to Fig. 1.

Assignment of Reference Numbers

A, B, C	-	characteristic of proteins
a, b, c	-	fraction

Patent Claims

1. Method for the multidimensional analysis of a proteome in which the biological material with the proteome to be analyzed is solubilized and the proteins belonging to the proteome are separated, quantitatively determined and identified, characterized in that the proteins of the proteome are subjected to a number n of different separating processes for n>2 under standardized conditions in such a way that each of the liquid fractions m_1 obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by τ different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that after combining the analysis data an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data space.

2. Method according to claim 1, characterized in that methods which separate according to the size of the protein and/or methods which separate according to the mass of the protein and/or methods which separate according to the charge of the protein and/or methods which separate according to the hydrophobicity of the protein and/or methods which separate according to the shape of the protein and/or methods which separate according to the affinity of the protein, with respect to specific ligands, also to antibodies are selected as separating methods.

3. Method according to claim 1, characterized in that methods for determining specific immunological characteristics and/or methods for determining specific catalytic activity and/or methods for determining chemical modification of the proteins of the proteome are used as identification methods.

4. Method according to claim 1, characterized in that methods for nonspecific determination of protein concentration with different sensitivities

and/or quantitative determination methods for determining specific catalytic activities and/or quantitative immunological methods and/or quantitative binding assays are selected as quantification methods.

5. Method according to claim 1, characterized in that the identification of individual proteins of the proteome is carried out directly by mass determination of the proteins.

6. Method according to claim 1, characterized in that the identification of individual proteins is carried out according to protease digestion and mass identification of fragments.

7. Method according to claim 1, characterized in that after the separation step the fractions are assembled in a two-dimensional multiple vessel system, preferably in the manner of and with the layout of microtitration plates.

8. Method according to claim 1, characterized in that in the first separating step the fractions are assembled in a defined grid, preferably in the n * 96 grid of microtitration technology.

9. Method according to claim 1, characterized in that all identification and quantification steps are carried out in a defined grid, preferably in the n * 96 grid, with adaptable liquid handling technique.

10. Method according to claim 9, characterized in that all identification steps and quantification steps are carried out with at least four two-dimensionally arranged, simultaneously working pipettor channels.

11. Method according to claim 1, characterized in that the first dimension for separation is high-resolution size exclusion, ion exchange or hydrophobicity chromatography, which are known per se, in that the second dimension is carried out by parallel separation and fractionation of the fractions of

the first dimension by means of a principle of separation other than that used for the first dimension, and in that each further separation and fractionation is carried out by parallel separating and fractionating methods with the fractions obtained from the preceding separating and fractionating steps.

12. Method according to claim 1, characterized in that the analysis data for the n-dimensional image of the protein are assembled in a database.

Abstract

The invention is directed to a method for the multidimensional analysis of a proteome. The method is used in biochemistry, biotechnology, medicine and in the pharmaceutical industry for purposes including diagnostics and the development of biologically active substances. It is the object of the invention to improve and facilitate quantification and identification of the proteins of a proteome and to make it possible for certain proteins to be quantified and identified for the first time. According to the invention, the proteins of the proteome are subjected to a number n of different separating processes under standardized conditions in such a way that each of the liquid fractions m_1 obtained in a separating step supplies m_2 liquid fractions in a subsequent separating step, wherein, after n separating steps, there are $m_1 * m_2 * \ldots m_n = M$ liquid fractions which are identified by o different analysis processes qualitatively and/or quantitatively by identification processes, known per se, and determined quantitatively by quantification processes which are likewise known per se, so that, after combining the analysis data in a database, an n-dimensional image of the proteome is obtained which is characterized by identifiers and quantifiers and by the position in the n-dimensional data network.

Prot. Nr.	Frak .a	Fra k.b	Fra k.c
404	1	1	9
218	1	3	3
970	1	4	1
777	1	4	10
325	1	5	1
190	1	5	7
348	1	6	2
459	1	6	4
156	1	8	6
172	1	8	7
999	1	8	10
598	1	10	9
189	1	10	9
281	2	1	1
618	2	2	10
274	2	4	10
849	2	5	2
494	2	6	2
81	2	7	3
564	2	10	2
670	2	10	4
562	3	1	8
805	3	2	3
108	3	2	9
923	3	3	3
386	3	3	5
295	3	4	5
36	3	7	1
20	3	10	4
606	3	10	5
592	4	1	3
8	4	2	6
917	4	2	8
756	4	3	8
255	4	4	8
483	4	4	9
503	4	6	2
991	4	6	9
741	4	9	8
507	4	10	7
69	5	1	4
584	5	2	4
26	5	2	10
771	5	3	4
515	5	3	10
188	5	5	6
566	5	6	5
478	5	7	9
529	5	8	3
149	5	8	3
627	5	8	7
729	5	9	4
956	5	9	8
832	5	10	4
897	5	10	5
792	5	10	10
350	6	1	6
360	6	3	2
423	6	3	9
875	6	4	1
82	6	5	2
578	6	5	8
487	6	7	6
363	6	8	1
333	6	9	9
373	6	10	4
699	7	1	1
508	7	1	8
480	7	1	9
582	7	2	10
531	7	2	10
177	7	6	9
918	7	7	4
239	8	1	6
649	8	1	9
106	8	2	1
610	8	2	6
738	8	5	3
161	8	5	8
154	8	7	5
981	8	8	5
972	9	3	5
96	9	3	7
192	9	4	4
694	9	6	2
691	9	6	2
952	9	6	8
68	9	8	1
913	9	8	9
950	9	9	5
150	9	10	6
137	10	1	9
80	10	2	1
969	10	4	1
778	10	5	3
462	10	7	8
737	10	8	8
785	10	9	9
802	10	10	3
856	11	1	1
467	11	6	8
932	11	9	4
927	11	10	1
326	11	10	8
594	11	10	9
523	12	2	7
119	12	3	1
114	12	3	1
297	12	4	5
447	12	4	7
238	12	8	3
605	12	8	5
558	12	8	8
169	12	9	7
43	12	9	7
987	12	10	3
191	12	10	9
866	13	1	6
997	13	2	6
852	13	2	9
142	13	3	5
613	13	4	6
338	13	5	4
690	13	5	6
398	13	5	6
669	13	6	1
903	13	7	1
656	13	7	8
537	13	8	10
779	13	9	8
708	14	1	6
141	14	2	8
236	14	3	3
724	14	4	7
151	14	6	7
275	14	7	1
304	14	7	6
726	14	7	10
664	14	10	6
214	15	1	7
525	15	1	10
263	15	2	4
181	15	2	10
289	15	3	5
731	15	5	3
547	15	6	3
877	15	6	10
631	15	7	6
596	15	7	7
202	15	7	10
535	15	9	5
193	15	9	5
800	16	1	7
187	16	1	8
560	16	3	3
98	16	5	4
714	16	6	1
680	16	7	3
652	16	8	7
454	16	8	8
390	16	8	8
876	16	9	5
369	16	9	8
116	16	10	2
770	17	2	4
902	17	3	6
988	17	3	7
957	17	4	4
183	17	4	10
946	17	5	6
402	17	5	6
823	17	5	7
593	17	9	1
977	18	3	8
221	18	6	2
498	18	6	4
752	18	6	7
727	18	6	7
996	18	8	3
438	18	8	7
521	18	9	9
968	18	10	2
735	18	10	10
90	19	1	4
443	19	1	7
361	19	2	10
3	19	3	8
436	19	3	10
662	19	6	4
79	19	8	3
273	19	9	6
403	19	9	8
493	19	10	1
524	19	10	2
914	19	10	4
246	20	1	1
272	20	3	1
985	20	6	2
841	20	6	3
634	20	6	5
540	20	6	5
385	20	8	5
316	20	10	10
70	21	1	3
470	21	2	5
66	21	2	5
904	21	3	8
696	21	3	8
795	21	4	4
23	21	6	5
811	21	8	8
29	21	9	5
362	21	10	3
709	22	1	6
347	22	2	8
695	22	3	1
663	22	4	4
597	22	4	7
579	23	2	6
614	23	2	10
434	23	3	9
22	23	8	3
420	23	8	5
983	23	10	5
975	23	10	6
244	24	1	1
688	24	2	2
561	24	3	9
257	24	4	4
5	24	5	5
280	24	6	8
212	24	8	7
896	25	2	2
389	25	2	5
815	25	3	8
331	25	4	5
589	25	6	8
814	25	6	9
816	25	7	2
67	25	7	4
845	25	7	10
890	25	9	5
104	25	9	8
491	25	10	4
568	26	1	8
379	26	2	6
909	26	3	8
687	26	4	8
105	26	5	1
858	26	6	5
827	26	6	8
798	26	7	10
769	27	1	5
976	27	3	1
415	27	3	6
296	27	3	10
921	27	4	1
99	27	4	6
851	27	4	8
240	27	5	9
519	27	8	2
766	27	8	5
51	27	8	7
451	27	8	9
581	27	9	6
288	27	9	6
647	28	2	10
242	28	4	8
757	28	5	7
450	28	6	3
707	28	9	2
611	28	9	10
426	28	9	10
463	29	1	6
259	29	2	4
19	29	4	3
530	29	4	6
72	29	4	6
60	29	5	1
603	29	5	2
907	29	6	5
660	29	7	4
685	29	8	7
411	29	10	4
829	29	10	5
797	30	3	3
383	30	3	6
600	30	3	8
472	30	4	9
32	30	5	2
542	30	6	3
953	30	6	6
391	30	8	9
171	30	10	4
123	30	10	8
224	31	1	2
418	31	2	6
964	31	2	10
773	31	3	7
762	31	4	4
159	31	5	3
73	31	5	8
409	31	7	9
783	31	8	7
370	31	9	5
874	31	10	4
733	32	1	10
100	32	2	4
502	32	2	9
381	32	3	1
786	32	3	10
638	32	4	5
859	32	4	6
599	32	5	8
448	32	7	3
321	32	7	3
871	32	7	7
118	32	9	6
622	32	9	10
607	32	10	7
84	33	1	3
83	33	3	1
939	33	3	2
925	33	4	7
16	33	5	3
653	33	5	4
208	33	6	3
567	33	7	2
194	33	8	1
58	33	8	2
499	33	8	10
198	33	8	10
700	33	9	3

Fig. 1a

869	33	10	3
526	34	3	3
332	34	3	4
636	34	4	6
121	34	4	8
998	34	5	1
355	34	5	8
346	34	5	10
270	34	6	1
810	34	6	2
34	34	6	9
734	34	7	1
862	34	7	6
164	34	9	4
157	34	10	8
796	35	1	6
962	35	1	8
736	35	3	4
85	35	3	8
47	35	4	2
793	35	4	6
819	35	6	6
671	35	6	8
432	35	6	10
195	35	9	1
324	35	9	4
658	35	10	3
468	36	2	1
643	36	4	2
926	36	5	8
693	36	8	7
767	36	9	3
354	36	10	7
955	37	1	1
314	37	4	4
548	37	5	8
313	37	6	10
219	37	7	4
959	37	8	5
46	37	9	2
497	38	1	2
678	38	3	5
260	38	3	8
754	38	3	9
648	38	4	7
310	38	6	4
209	38	6	4
990	38	6	7
11	38	8	10
941	38	9	5
184	38	9	8
637	39	1	8
545	39	2	7
163	39	7	9
267	39	8	8
1	39	9	6
229	39	10	1
53	39	10	4
822	40	1	3
891	40	1	7
335	40	3	6
623	40	4	2
901	40	5	8
828	40	6	6
666	40	7	10
474	40	10	2
704	40	10	3

979	41	3	5
556	41	4	9
44	41	5	3
812	41	7	8
37	41	9	6
343	41	10	8
911	42	1	6
522	42	1	10
505	42	3	1
417	42	3	2
782	42	4	3
807	42	5	7
765	42	6	3
168	42	8	1
857	42	8	9
657	42	9	1
252	42	10	1
475	42	10	3
173	42	10	4
302	43	2	3
809	43	2	10
431	43	4	3
906	43	5	10
602	43	7	6
283	43	7	6
492	43	9	4
349	43	9	7
364	43	10	7
197	43	10	9
465	44	2	1
549	44	2	4
635	44	3	9
538	44	5	6
801	44	6	6
993	44	6	7
965	44	7	6
780	44	8	2
830	44	9	2
277	44	9	5
269	44	9	7
113	45	1	1
265	45	1	3
710	45	1	6
477	45	2	1
922	45	2	6
668	45	4	2
271	45	4	7
863	45	6	2
39	45	7	5
948	45	7	9
376	46	1	5
428	46	5	5
317	46	8	1
117	46	9	6
689	47	1	5
992	47	2	4
559	47	2	6
375	47	3	10
554	47	5	4
624	47	6	6
565	47	6	6
457	47	9	3
17	47	10	4
864	48	1	1
372	48	2	4
759	48	2	7
129	48	3	4

938	48	5	10
511	48	6	4
485	48	6	4
915	48	7	5
458	48	7	8
138	48	7	8
120	48	7	10
868	48	8	3
486	48	10	10
30	49	1	6
571	49	2	1
936	49	3	10
520	49	6	10
775	49	8	4
421	49	8	6
287	50	1	1
89	50	3	3
847	50	3	7
49	50	5	8
577	50	6	7
2	50	7	5
374	50	7	7
711	50	8	9
722	50	9	1
958	50	10	4
645	51	1	6
720	51	2	1
300	51	2	3
973	51	6	4
282	51	6	10
674	51	7	7
213	51	9	4
833	51	9	10
216	51	10	4
986	52	1	3
253	52	2	8
625	52	2	9
768	52	3	6
818	52	3	7
804	52	6	1
824	52	6	5
705	52	6	7
512	52	6	7
337	52	8	3
639	52	9	7
204	52	9	10
284	52	10	7
615	53	1	3
261	53	1	8
612	53	6	6
604	53	10	3
15	53	10	7
441	53	10	8
843	54	1	3
97	54	1	6
235	54	4	1
712	54	6	10
583	54	6	10
91	54	7	3
249	54	7	4
24	54	8	2
576	54	8	10
160	55	1	10
410	55	2	1
929	55	2	4
387	55	2	5
74	55	2	9

743	55	5	7
305	55	6	8
473	55	8	1
266	55	8	4
393	55	8	10
320	55	9	1
276	55	9	6
7	55	9	10
799	55	10	9
518	56	1	10
245	56	2	6
870	56	4	5
940	56	5	3
673	56	5	6
182	56	5	10
167	56	5	10
679	56	8	1
78	56	8	4
437	56	8	8
873	56	9	6
888	56	10	5
201	57	1	7
412	57	5	7
133	57	5	9
908	57	6	2
967	57	6	3
12	57	7	7
677	58	1	5
139	58	1	5
352	58	1	9
293	58	2	10
543	58	9	3
954	58	10	2
681	59	2	1
844	59	2	9
753	59	3	2
881	59	5	2
52	59	5	3
501	59	5	8
516	59	6	3
196	59	6	7
860	59	7	7
628	59	8	10
162	59	10	1
54	59	10	8
291	60	2	5
553	60	2	7
655	60	2	9
227	60	5	4
61	60	5	4
165	60	8	1
199	60	8	2
210	60	10	2
442	60	10	6
728	60	10	9
633	61	1	1
570	61	1	4
220	61	1	6
179	61	5	5
808	61	7	8
223	61	9	2
336	62	1	10
232	62	2	6
77	62	2	8
399	62	2	10
55	62	3	8
675	62	3	10

834	62	4	2
471	62	4	3
919	62	4	10
342	62	7	3
460	62	7	8
378	62	7	10
94	62	10	7
935	63	1	7
395	63	2	4
464	63	2	8
949	63	3	2
394	63	4	1
14	63	4	6
683	63	6	4
298	63	6	10
698	63	7	5
920	63	7	9
481	63	7	9
817	63	8	8
76	63	8	9
416	63	8	10
371	63	9	3
739	63	9	5
646	63	9	6
135	63	9	7
233	63	9	10
237	64	1	4
764	64	4	3
974	64	5	6
62	64	5	6
745	64	6	5
248	64	6	6
585	64	6	9
466	64	9	2
217	64	9	5
730	64	9	8
761	64	10	6
569	64	10	8
750	65	1	3
38	65	2	4
102	65	3	8
880	65	4	8
528	65	5	1
725	65	7	10
787	65	8	1
533	65	9	1
408	65	9	8
882	66	1	2
264	66	1	5
87	66	1	6
429	66	2	5
867	66	3	5
539	66	3	5
789	66	3	8
328	66	4	9
960	66	6	3
640	66	8	8
794	66	9	1
452	67	1	5
500	67	1	8
510	67	3	6
10	67	5	2
650	67	6	3
490	67	6	9
586	67	7	9
344	67	8	4
482	67	8	7

Fig. 1b

651	68	1	4
368	68	3	8
703	68	4	5
226	68	4	5
33	68	4	5
706	68	7	5
158	68	7	8
178	68	9	3
788	68	9	10
145	69	1	4
250	69	3	10
131	69	5	6
425	69	6	8
854	69	8	3
303	69	8	5
899	69	9	9
186	69	10	1
721	69	10	4
455	70	1	2
125	70	1	7
122	70	2	6
256	70	4	1
928	70	4	2
842	70	4	4
484	70	4	5
308	70	4	8
222	70	5	8
641	70	6	3
740	70	6	4
56	70	7	3
620	70	7	9
101	70	8	2
86	70	8	2
848	70	8	8
595	70	9	2
898	70	9	9
716	70	9	10
329	70	10	9
885	71	2	8
18	71	2	9
676	71	3	3
702	71	3	8
128	71	7	4
140	71	7	6
13	71	8	2
621	71	8	3
642	71	8	5
334	71	9	4
489	71	10	2
742	72	3	1
632	72	3	4
247	72	3	8
506	72	4	3
732	72	4	9
853	72	8	3
392	72	9	3
92	72	9	5
552	73	2	3
723	73	4	3
994	73	4	4
945	73	6	8
665	73	6	8
414	73	7	8
879	73	10	2
661	74	2	2
984	74	3	2
905	74	5	1

887	74	5	8
461	74	6	7
546	74	6	8
575	74	7	9
45	74	7	10
290	75	1	4
846	75	2	1
440	75	6	7
155	75	7	10
31	75	8	5
9	75	9	8
127	75	10	4
748	76	1	8
405	76	2	9
555	76	3	6
180	76	5	8
630	76	10	3
312	76	10	7
718	77	1	9
889	77	2	8
772	77	3	5
444	77	3	5
380	77	4	9
64	77	5	10
982	77	8	5
825	77	10	3
840	78	2	1
306	78	2	5
865	78	6	4
170	78	7	3
839	78	8	7
59	79	1	3
644	79	4	4
307	79	4	9
327	79	5	5
251	79	5	6
544	79	6	2
230	79	6	4
341	79	7	7
763	79	8	2
132	79	9	1
629	79	9	6
406	80	1	2
850	80	3	1
35	80	3	3
279	80	5	5
884	80	5	8
572	80	5	10
947	80	6	3
427	80	10	9
446	81	1	3
616	81	1	7
175	81	2	2
654	81	2	4
205	81	4	8
206	81	4	10
872	81	7	2
93	81	7	6
791	81	8	4
469	81	8	5
262	82	2	7
514	82	3	1
776	82	3	8
377	82	3	10
203	82	4	7
148	82	5	10
1000	82	6	5

900	82	6	8
413	82	6	10
110	82	8	5
551	83	1	2
439	83	1	9
445	83	3	1
353	83	3	2
294	83	3	3
883	83	3	10
580	83	7	2
715	83	8	3
4	83	8	4
931	83	8	6
590	83	10	5
323	83	10	8
63	84	5	2
989	84	6	5
152	84	8	9
207	84	9	2
806	84	10	2
111	84	10	7
513	85	1	2
527	85	1	10
109	85	2	7
95	85	3	8
315	85	3	9
130	85	4	6
241	85	5	5
601	85	7	3
504	85	8	4
944	85	10	9
755	86	1	1
534	86	1	5
147	86	3	2
146	86	3	2
751	86	3	5
667	86	3	8
48	86	5	2
258	86	5	6
107	86	6	8
358	86	8	2
820	86	8	4
254	86	9	1
971	86	9	4
496	86	9	8
701	86	10	7
659	87	1	5
943	87	1	8
894	87	2	5
517	87	3	3
176	87	5	5
322	87	5	10
878	87	6	4
243	87	7	8
821	87	7	9
836	87	8	2
942	87	9	6
174	87	10	2
937	88	1	3
626	88	5	1
684	88	6	8
609	88	7	2
826	88	9	5
88	89	2	1
330	89	3	5
790	89	4	1
388	89	4	8

588	89	6	5
228	89	6	6
838	89	7	5
456	89	8	2
435	89	8	7
479	89	10	5
980	90	1	2
424	90	1	3
837	90	4	5
143	90	4	10
57	90	5	2
144	90	6	6
430	90	7	7
319	90	8	2
400	90	8	4
951	90	8	7
719	90	9	10
345	90	10	3
366	90	10	8
749	90	10	10
339	91	1	3
65	91	1	3
831	91	2	2
278	91	2	7
136	91	2	10
453	91	3	2
407	91	8	3
682	91	8	4
886	91	8	7
781	91	9	5
672	91	10	4
384	92	1	2
855	92	2	5
75	92	4	3
285	92	4	10
166	92	5	1
356	92	5	9
6	92	6	2
299	92	6	6
185	92	6	9
758	92	7	4
495	92	10	2
912	92	10	6
359	93	2	6
747	93	2	9
40	93	5	2
301	93	5	3
433	93	5	7
861	93	6	6
234	93	7	1
541	93	8	1
200	93	10	1
587	93	10	8
619	94	1	5
449	94	1	5
134	94	1	9
357	94	1	10
933	94	2	2
617	94	2	3
963	94	2	10
103	94	3	4
41	94	3	7
25	94	8	3
608	94	9	4
934	95	1	7
318	95	1	7
550	95	2	4

574	95	3	5
27	95	6	6
717	95	6	7
71	95	9	2
995	95	9	3
744	95	9	4
924	95	9	8
563	96	1	5
895	96	4	7
50	96	5	7
813	96	5	9
916	96	7	3
893	96	7	4
686	96	7	8
365	96	7	10
231	96	8	5
21	96	9	4
713	96	9	6
367	96	10	4
351	97	2	9
774	97	3	4
476	97	5	1
422	97	5	10
382	97	7	1
124	97	8	2
961	97	8	5
697	97	9	5
225	97	10	3
112	97	10	5
760	98	2	1
536	98	2	8
488	98	3	7
692	98	5	1
401	98	5	2
591	98	5	4
153	98	5	10
292	98	7	1
286	98	7	2
396	98	7	4
509	98	8	8
397	98	9	6
557	98	10	4
978	99	1	3
803	99	4	7
892	99	4	8
746	99	5	7
215	99	7	1
28	99	7	4
930	99	8	5
340	99	9	4
835	99	10	9
526	99	10	9
311	100	2	6
268	100	2	6
211	100	2	10
910	100	4	2
309	100	4	6
419	100	5	2
784	100	8	1
42	100	8	6
115	100	8	8
966	100	9	8
573	100	9	9
532	100	10	5

Fig. 1c

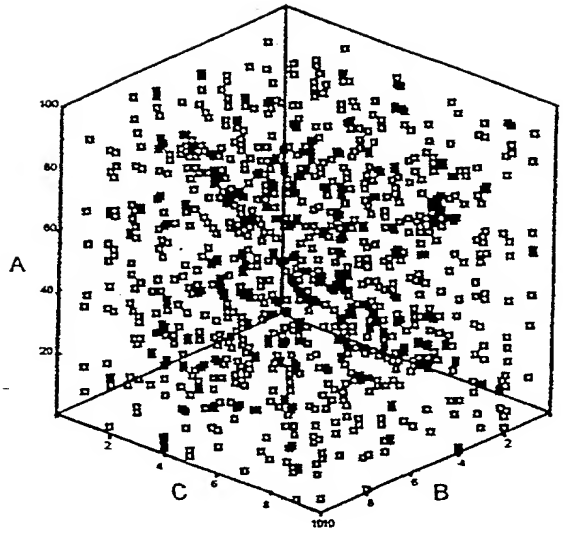

Fig. 2

UNITED STATES OF AMERICA COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION	FILE NO. GK-OEH-120/ 500814.20021

As a below named inventor, I hereby declare that: my residence, post office address and citizenship are as stated below next to my name; that I verily believe that I am the original, first and sole inventor (if only one name is listed below) or a joint inventor (if plural inventors are named) of the subject matter which is claimed and for which a patent is sought on the invention entitled:

METHOD FOR THE MULTI-DIMENSIONAL ANALYSIS OF A PROTEOME

The specification of which

- ☐ is attached hereto.
- ☐ was filed on _____ as United States patent application Serial Number _____.
- ☒ was filed on July 4, 2000 as PCT international patent application No. PCT/DE00/02154 and was amended on _____ (if any).

I hereby state that I have reviewed and understand the contents of the above identified specification, including the claims, as amended by any amendment referred to above.

I acknowledge the duty to disclose all information known to be material to patentability in accordance with Title 37, Code of Federal Regulations, § 1.56.

I hereby claim foreign priority benefits under Title 35, United States Code § 119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application on which priority is claimed:

Prior Foreign Application(s)

COUNTRY	APPLICATION NUMBER	DATE OF FILING *(day, month, year)*	PRIORITY CLAIMED UNDER 35 U.S.C. § 119
Germany	199 32 270.8	05 July 1999	YES x NO ____
			YES ____ NO ____

I hereby appoint REED SMITH LLP and the members of the firm: Lloyd McAulay, Reg. No. 20,423; J. Harold Nissen, Reg. No. 17,283; Jules E. Goldberg, Reg. No. 24,408; Gerald H. Kiel, Reg. No. 25,116; Eugene LeDonne, Reg. No. 35,930; Stephen Chin, Reg. No. 39,938; Arthur Dresner, Reg. No. 24,403; Daniel Lent, Reg. No. 44,867; Samir R. Patel, Reg. No. 44,998; and Harry K. Ahn, Reg. No. 40,243, as attorneys with full power of substitution and revocation to prosecute all business in the Patent & Trademark Office connected therewith and to receive all correspondence.

SEND CORRESPONDENCE TO: Gerald H. Kiel, Esq.
REED SMITH , LLP
375 Park Avenue
New York, New York 10152-1799, U.S.A.

DIRECT TELEPHONE CALLS TO: (212) 521-5400

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

FULL NAME OF SOLE OR FIRST INVENTOR	INVENTOR'S SIGNATURE	DATE
Thomas MOORE		08.04.02
RESIDENCE D-07751 Drackendorf, Germany DEX		COUNTRY OF CITIZENSHIP Germany
POST OFFICE ADDRESS Zur Laemmerlaide 15, D-07751 Drackendorf, Germany		
FULL NAME OF SECOND INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
Anton HORN		08.04.02
RESIDENCE D-07749 Jena, Germany DEX		COUNTRY OF CITIZENSHIP Germany
POST OFFICE ADDRESS Pennickental 4, D-07749 Jena, Germany		

☒

COMBINED DECLARATION AND POWER OF ATTORNEY FOR PATENT APPLICATION (continued) 3-8D		File No. GK-OEH-119/ 500814.20021
FULL NAME OF THIRD INVENTOR (IF ANY) Stefan KREUSCH	INVENTOR'S SIGNATURE S. Kreusch	DATE 08.04.02
RESIDENCE D-07743 Jena, Germany DE		COUNTRY OF CITIZENSHIP Germany
POST OFFICE ADDRESS Freiligrathstrasse 90, D-07743 Jena, Germany		
FULL NAME OF FOURTH JOINT INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		
FULL NAME OF FIFTH INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		
FULL NAME OF SIXTH JOINT INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		
FULL NAME OF SEVENTH JOINT INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		
FULL NAME OF EIGHTH JOINT INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		
FULL NAME OF NINTH JOINT INVENTOR (IF ANY)	INVENTOR'S SIGNATURE	DATE
RESIDENCE		COUNTRY OF CITIZENSHIP
POST OFFICE ADDRESS		